Relationship between demographic and genetic population size and role of the environment in a stream fish

Thaïs A. Bernos

A thesis
in
The Department
Of
Biology

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science (Biology) at
Concordia University
Montreal, Quebec, Canada

February 2016

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Thaïs A. Bernos

Entitled: Relationship between demographic and genetic population size and role of the environment in a stream fish

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Biology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____ Chair
Dr. Robert Weladji

_____ Examiner
Dr. Selvadurai Dayanandan

_____ Examiner
Dr. David Welsh

_____ Examiner
Dr. James Grant

_____ Supervisor
Dr. Dylan J. Fraser

Approved by _____
Dr. Selvadurai Dayanandan, Graduate Program Director

_____
Dr. André Roy, Dean of Faculty

Date _____

ABSTRACT

Relationship between demographic and genetic population size and role of the environment in a stream fish

Thaïs A. Bernos

As wild populations become increasingly small and vulnerable, conservation managers typically must make quick decisions based on limited resources. Two crucial parameters affecting management decisions are the census population size (N) and the effective number of breeders (N_b). However, measuring N and N_b is often difficult, making it of interest to generalize one from the other. We assessed the relationship between N and N_b from 2008-2015 in twelve brook trout populations varying greatly in N (49-10032) and N_b (3-567). Most of the variability in N_b could be explained by N (R^2_m=0.54, p<0.001) or stream length (R^2_m=0.44, p<0.001) alone. The ratio N_b/N increased at small N or following an annual decrease in N (R^2=0.49, p<0.01), suggesting density-dependent constraints on N_b/N (genetic compensation). We did not find any evidence for consistent differences invariability in N_b and/or N_b/N between small and large populations; however, small populations had more varying temporal variability in N_b/N ratios than large populations. N_b and N_b/N were respectively 2.5-fold and 2.3-fold more variable among populations than temporally within populations. Collectively, our results suggest that conservation resources could be saved by using N or N_b to infer the other to assess relative population sizes. However, using one variable to infer the other to monitor trends within populations is less recommended, perhaps even less so in small populations given their less predictable N_b vs. N dynamics.

Acknowledgements

We thank M. Yates, S. Smart, J. Wood, P. Debes, K. Marin, M. Heath, Z. Wells, C. Zastavniouk, E. Lawrence, J. Strong and R Carson for fieldwork or laboratory assistance, the Mistaken Point Ecological Reserve and the Department of Fisheries and Oceans, Newfoundland, for allowing this work. We also thank M. Yates and K. Marin for statistical advice and helpful reviews of this manuscript. This research was supported by NSERC Discovery and Accelerator Grants to D. Fraser, as well as two excellence awards from the Quebec Center for Biodiversity Science and a Vladimir J. Elgart Graduate Entrance Scholarship to T. Bernos.

Contributions of authors

T.A. Bernos conducted fieldwork and molecular work, statistical analyses, and redaction of the manuscript. D.J. Fraser provided field and logistical support, theoretical guidance and writing assistance

Table of contents

List of Tables

List of Figures

Introduction

Conservation biologists and population geneticists recognize the importance of the census population size (N) and effective number of breeders (N_b) for conservation. While demographic events contribute to changes in N (recruitment failure, increased mortality, etc.), genetic processes (drift, inbreeding) are influenced by N_b. N_b also provides valuable information about a population's reproductive biology and is often used as a proxy for the effective population size (N_e) in iteroparous species (Waples 2002; Waples *et al.* 2013, 2014; Duong *et al.* 2013). Although N and N_b provide information on the demographic and genetic status of small and endangered populations, both parameters can be challenging and resource intensive to estimate in practice.

Several recent publications suggested that N_b could be used to infer N (or vice versa) provided that N_b/N ratios remain stable across time or within specific taxa (Schwartz *et al.* 2007; Tallmon *et al.* 2010; Palstra & Fraser 2012). Nevertheless, while some studies found positive correlations between N and N_b (Osborne *et al.* 2010; Charlier *et al.* 2012; Dowling *et al.* 2014), others found no relationship at all (Ardren & Kapuscinski 2003; Araki *et al.* 2007; Palstra *et al.* 2009; Berry & Kirkwood 2010; Johnstone *et al.* 2012; Serbezov *et al.* 2012; Belmar-Lucero *et al.* 2012; Duong *et al.* 2013; Whiteley *et al.* 2015). These disparities could be partly attributed to a number of factors; for instance, some studies either included only a small number of populations, had no temporal replication, incorrectly linked N_b and N, and/or could not estimate N with precision for logistical reasons.

A first step in understanding the link between demographic population size, genetic population size and the environment is to explore spatial variation in N_b and N_b/N ratios across a broad range of populations of the same species. It is necessary to consider populations varying in size, as previous studies suggested that changes in ecological interactions at smaller population sizes can inflate N_b/N in various species, including salmonid fishes(Ardren & Kapuscinski 2003; Araki *et al.* 2007; Fraser *et al.* 2007; Palstra & Ruzzante 2008; Perrier *et al.* 2015). In salmonids, genetic compensation could arise because the competition for access to mates or spawning grounds could be relaxed at low N (Chebanov 1991) which in turn might reduce the variance in

reproductive success known to drive N_b below N. Apart from one notable exception (Whiteley *et al.* 2015), this has led most researchers to use correlations between the N_b/N ratio and N as a means to infer the occurrence of genetic compensation. However, relating a ratio to its denominator (N_b/N to N) results in spurious statistical relationships (Jackson & Somers 1991). Consequently, how N_b/N varies across populations of various sizes remains uncertain.

It is also important to assess N_b, N, and N_b/N across sequential years within populations because they are likely affected by specific environmental variables (Kanno *et al.* 2015) and/or fluctuating environmental conditions. For example, habitat conditions can become more variable as populations become smaller and fragmented, thus increasing the variability in environmental selective pressures (Fraser *et al.* 2014; Wood *et al.* 2014). In some years, some small populations might therefore experience ecological conditions driving N_b below N by reducing recruitment, while other small populations might be found in higher quality, stable habitats that are simply too small to accommodate larger populations. Alternatively, large populations might be found in larger habitats that are more spatially heterogeneous in environmental conditions and may therefore be more stable.

Salmonid species such as brook trout (*Salvelinus fontinalis*) are ideal models for investigating the relationship between N, N_b, N_b/N ratios, and the environment. Indeed, certain aspects of their reproductive biology and environmental features are likely to result in different N_b/N ratios among populations. For instance, spawning habitat requirements, such as groundwater upwelling, can limit the availability of suitable spawning sites for brook trout, thus resulting in the disproportionate contribution of some individuals to the next generation, a potential mechanism for genetic compensation (Blanchfield and Ridgway 1997, 2005; Curry and Neakes 1995; Guillemette *et al.* 2011). Additionally, brook trout exhibit considerable life history and phenotypic variation in relation to local habitat features (Hutchings 1993; Fraser & Bernatchez 2005). This is likely to affect intra-specific competition for access to spawning resources, for example through density-dependent effects. Recently, Wood *et al.* (2014) and Whiteley *et al.* (2013, 2015) found striking correlations between N_b and either habitat size or spawning habitat quantity or quality among different brook trout populations. Collectively, these points raise the possibility that with only a minimum amount of information about a population or its

environment, one might be able to derive accurate estimates of genetic and/or demographic population sizes from the other.

There is a need to characterize the relationship between N and N_b across a wide range of population sizes and spatiotemporal environmental conditions. To fill this gap, we monitored twelve brook trout populations between 2008-2015 to investigate the effect of environmental and population size changes on N_b and N_b/N variance, both among and temporally within populations. Among populations, we tested for 1) a positive correlation between N_b and N and 2) a significant correlation between specific environmental variables, N_b and N. Within populations, we predicted 3) a positive correlation between annual changes in N and N_b and 4) a negative correlation between annual changes in N and N_b/N as predicted by the genetic compensation hypothesis. We also predicted 5) greater variability in N_b and N_b/N among than temporally within populations; 6) a negative correlation between population size and temporal variability in N_b and N_b/N, and; 7) greater temporal variability in N_b and N_b/N among small populations based on previous work (Wood *et al.* 2014). Our study provides valuable insights into how demographic and genetic factors interact in populations of varying size. It also has relevance in terms of effective population monitoring of small endangered populations, large populations harvested by humans, or populations where obtaining demographic data is more challenging (Palstra & Fraser 2012).

Materials and methods

Study site

Our twelve CR brook trout populations originate from a common ancestor (Danzmann *et al.* 1998) and inhabit a fine geographic scale (400km^2). They vary widely in population size (Wood et al. 2014), phenotypic characteristics (Wood et al. 2015), life history (Hutchings 1993; Belmar-Lucero *et al.* 2012) and putative adaptive genetic differentiation (Fraser *et al.* 2014). Their stream habitats also differ greatly in environmental characteristics that affect brook trout reproductive biology, including the number, area, and quality of spawning sites (Belmar-Lucero *et al.* 2012). Additionally, the ecology and habitat of CR brook trout can be comprehensively sampled since they inhabit small streams (0.27-8.10km in length). Finally, they do not experience significant interspecific competition or predation: only three of the streams (UO, LO, and WN) contain small, natural populations of Atlantic salmon (*Salmo salar)*; two additional streams (CC, WN) harbor small threespine stickleback populations (*Gasterosteus aculeatus)* and occasional American eels (*Anguilla rostrata)*. Collectively, the diversity among CR trout populations offers an ideal model for the study of the relative role of environmental and demographic factors in shaping N_b/N ratios.

Adult census population size (*N*)

We estimated *N* each summer (June-July) for 63 cohorts across 12 populations from 2010 to 2015. *N* was defined as the total number of adults; this corresponds to trout of age 1+ and older at CR (Hutchings 1993; Belmar-Lucero *et al.* 2012). The size distribution of age-0 and age-1+ trout does not overlap. The sampling protocol in 2012-2015 was standardized accordingly to the years 2010 and 2011 described in detail in Wood et al. (2014). *N* was estimated from traditional mark-recapture methods with a single recapture event (Petersen 1985). Individuals were marked by adipose fin clips or numerical floy tags (FD-68B Fine Fabric Anchor Tags; Floy Tag and Manufacturing, LOC), depending on the population. Recapture events took place approximatively one week following tagging.

Genetic data

We sampled a total of 71 young-of-the-year (YOY or age 0) cohorts across 12 populations from 2010 and 2015 to derive cohort N_b estimates within populations corresponding to the years 2008 to 2014. Caudal fin clips were preserved in 95% ethanol until DNA was extracted using a phenol-chloroform protocol in sampling years 2010-2012 (described in Belmar-Lucero *et al.* 2012) and a chelex protocol in 2013-2015 (adapted from Hua and Orban 2005). YOY were sampled evenly along each stream to maximize both the potential number of families sampled and the sample size for precise N_b estimates (see Table 1 for sample sizes). In total, 4796 YOY genotypes were included in this study, 1306 of which originated from Wood *et al.* (2014; years 2010/2011). All trout were genotyped at 13 microsatellite loci; details of PCR conditions, electrophoresis using a Genetic Analyzer and allele scoring are found in Wood *et al.* (2014); alleles were manually scored using Genemapper 3.2 and Peak scanner (Applied Biosystems Inc.). As two study loci (*Sco204*, *Sco218*) were found to be linked in a considerable number of cohort samples as these were further developed through time, *Sco218* was removed from subsequent analyses.

Population genetic analyses

Departures from linkage and Hardy Weinberg equilibrium were verified using GENEPOP V4.2 across the 71 sampled YOY cohorts (Raymond & Rousset 1995). Statistical significance levels were adjusted to control for type I error via Bonferroni correction. We used GenALEx V6.5 (Peakall & Smouse 2012) to quantify alleles per locus as well as observed and expected heterozygosity for each locus in each cohort sample. Although most stream drainages at CR constitute single populations of brook trout (Wood et al. 2014), we also used STRUCTURE V2.3.4 (Pritchard *et al.* 2000) to evaluate the temporal stability of within drainage population structuring found in 2010/2011 in four drainages by Wood et al. (2014). Four independent runs with a burn-in period of 100000 followed by 150000 iterations per drainage were run under a model of admixture and correlated allele frequencies using K subpopulation values of 1- 5 and replicated 5 times per K to estimate posterior probabilities (ln P(D)) of the data. We determined K using a combination of the ΔK procedure of Evanno, Regnaut, and Goudet (2005), by interpreting the ln P(D) values themselves, and by assessing the strength of individual assignments within clusters. To avoid erroneous population structure inferences in STRUCTURE runs due to family groups (Anderson & Dunham 2008), we identified any full

sibling groups in cohort samples with COLONY V2.0.5.9 (Wang 2004) and removed all but one randomly selected individual in each such group inferred with a probability of 95% or higher. In three of four drainages, multiple populations were consistently demarcated across sampling years (see appendix A), and were associated with clear geographic divisions such as waterfalls (UC vs. LC) or fragmentation and isolation of stream beds (DY vs. UO vs. LO; BF vs. WN). A suggestion of subpopulation structure in the two streams of the FW drainage (FW vs. PN) was raised from both a higher posterior probability and delta K for K = 3 than for K = 2, and several heterozygote deficiencies (see Results). The lack of distinct spatial structuring between the three subpopulations suggested the existence of a cryptic population; the FW drainage was therefore excluded from subsequent analyses since it violated an important assumption of the LDNe method and it was impossible to estimate *N* in the different subpopulations.

Effective number of breeders (N_b)

N_b estimates were derived using LDNe (Waples and Do 2010), a linkage disequilibrium one sample estimator. LDNe uses the assumption that as N_b decreases, non-random associations between alleles increase due to allelic drift. To avoid upward bias due to rare alleles, we used an allele exclusion criterion P_{crit}=0.05 for sample sizes < 25, P_{crit}=0.02 for sample sizes between 25 and 100, and P_{crit}=0.01 for sample sizes exceeding 100 (Waples and Do 2010).

We obtained a total of 63 N_b estimates since we were unable to derive N_b (no lower confidence intervals and/or point estimates) from 8 of the 71 sampled cohorts. Of the 63 N_b estimates, 14 had upper confidence intervals including infinity; this lack of precision can be chiefly explained by small sample sizes despite intensive sampling efforts (LC, UC, LO, DY, HM) and/or a lack of genetic diversity (HM, STBC) in small or very small populations. For one additional population (BC), failure to converge to a reliable N_b in 2009 and 2011 was probably due to an interaction between sample size and a large N_b (see Waples & Do 2010).

Stream environmental data

To obtain fine-scale habitat coverage across the entire length of each stream, we collected habitat data annually from a total of 1,850 stream cross sections (Appendix A), with 18-32 cross-sections per population in 2010 and 18-64 cross-sections per population in subsequent years.

GPS coordinates of each cross-section were recorded in 2010 or 2011; the same cross-sections were sampled each year (+/- 0-5m due to fine scale GPS uncertainties). In the present study, we tested the influence of stream flow, temperature and stream length on N_b since previous work supported that these are important environmental variables shaping brook trout population dynamics (Whiteley et al. 2013, 2015; Wood et al. 2014; Xu, Letcher, and Nislow 2010). Flow (ms^{-1}) represented a surrogate for discharge (m^3s^{-1}), which itself was too strongly correlated with stream length to be included in our models (longer streams tend to be wider, deeper, and flow at higher velocity). Variations in flow could affect adult survival and body condition by reducing summer habitat availability and quality (Hakala & Hartman 2004; Xu *et al.* 2010). Additionally, high summer temperature can also affect survival and body condition (Xu *et al.* 2010); temperature-induced stress and metabolism changes can also affect gonad development (Pankhurst & Munday 2011), potentially leading to delayed spawning and reduced red construction (Warren et al. 2012). Stream flow was measured using a standard float technique (whiffle ball with a 1 meter string) once per summer to the nearest 0.01s; 3 float trails were evenly distributed across the stream width at each cross-section (Appendix A). Stream temperature was recorded every 75min between the months of June to September to the nearest 0.01*C using Maxim's ibutton temperature loggers. Loggers (two per stream) were affixed to rocks at the stream bottom (Appendix A). Stream length was measured as in Wood et al. (2014). The environmental variables were measured the same years as N, and they were therefore likely to affect the breeding adult generation described by N_b.

Relationship between N_b, N and the environment among populations

While a strong relationship between N and N_b would suggest that a simple conversion might exist between the two variables, previous evaluations of the relationship between N_b and N did not consider sequential years across a wide range of population sizes. We evaluated the relationship between N, N_b and the environment to test the predictions that there would exist (i) a positive correlation between N and N_b given that as the number of adults increases, more individuals may contribute to the next generation; (ii) a positive correlation between stream length and N and/or N_b as suggested by recent publications (Whiteley et al. 2013, 2015; Wood et al. 2014) and, (iii) a positive correlation between flow and N_b and a negative correlation between summer temperature and N_b (Xu *et al.* 2010; Letcher *et al.* 2014; Whiteley *et al.* 2015).

We used linear mixed models (LMM) to evaluate the effect of *N* and environmental variables on N_b. Environmental variables from all rivers in each sampled year were summarized as means, and *N* and N_b were natural log-transformed to increase normality. In all models, explanatory variables were centered and scaled. We included a random intercept-by-population term in all models, regardless of its significance, to account for basal variations in mean N_b.

The mixed model analyses were performed using the lme4 package (Pinheiro *et al.* 2013) in R Studio 0.99.484 (Team 2013). Model selection was performed on maximum likelihood fitted models. Final parameters estimates, marginal R^2 (Nakagawa & Schielzeth 2013; R^2_m: variability explained by fixed effects) and conditional R^2 (R^2_c: variability explained by fixed and random effects) were obtained using restricted maximum likelihood. Backward model selection was conducted by stepwise removal of non-significant fixed effects ($p > 0.05$) using Wald F tests with with denominator degrees of freedom obtained using the Kenward-Roger approximation (Kenward & Roger 1997) in the R package lsmeans (Lenth & Herva 2015).

Variance inflation factor (VIF) analyses were conducted in all models to check for multicollinearity between environmental variables and *N*. High VIF scores indicated that stream size and *N* were strongly collinear. To determine whether *N* or stream size best explained N_b, mixed models were fitted separately with *N* and stream size and the models' explained variance were compared. *N*, stream length, summer flow, and summer temperature were treated as fixed effects. To determine whether *N* or stream length alone explained a substantial amount of variability in N_b, reduced mixed models including only *N* or stream length as a fixed effect were fitted and the explained variances were compared with those of the fuller models. We did not include interaction terms in any of the models as it led to convergence failure.

We also used LMM to determine whether a significant correlation existed between *N* and environmental variables. The LMM included *N* as a response variable, stream length, summer flow and temperature as explanatory variable and a random intercept-by-population term. Stream length and *N* were natural log transformed to increase normality and reduce heteroscedasticity.

Relative trends in N, N_b and N_b/N within populations

A relationship between annual changes in N and N_b or N_b/N within populations would suggest that conservation resource use could be optimized by using only one variable to monitor genetic *and* demographic trends within natural populations. Two alternative predictions were investigated: (i) a positive correlation would exist between a change in N_b and a change in N, as more adults in the population could allow more parents to contribute; and, (ii) a negative correlation between a change in N and in N_b/N might be predicted. For the latter, according to the genetic compensation hypothesis, a substantial increase in N could heighten the competition for access to spawning areas, which would increase the variance in reproductive success and reduce N_b/N (Ardren & Kapuscinski 2003).

We used percent change (PC) to express relative annual changes in N, N_b and N_b/N where $PC=(V_x-V_{x-1})/V_{x-1}*100$ (V_x and V_{x-1} = value in year x and in the preceding year). As preliminary analyses suggested that our data fulfilled linear regression assumptions, we used PC in N_b or N_b/N (separately) as response variables and PC in N as the explanatory variable. Because this analysis required two consecutive N_b estimates correctly matched to two consecutive N estimates within populations, there were only a total of 24 PC data points for the analysis from our time series. We could not include by-population random intercepts to estimate within-population variability, as 3 of 12 populations only had one data point.

Patterns of spatial and temporal variability in N_b, N and N_b/N

As an exploratory analysis, we also tested a series of alternative predictions regarding the relative extant of spatial and temporal variability in N_b, N and N_b/N. This analysis is relevant towards a better understanding of the sampling conditions under which one variable can reliably be used to infer the other. First, we tested whether there was a significant difference in temporal variability between N_b and N. Indeed, Whiteley et al. (2015) suggested that the use of N_b to monitor trends was limited since it may be more temporally constrained than that of N. Second, we tested two predictions related to variability and N_b or N_b/N and population size: 1) small populations might be more temporally variable in N_b and/or N_b/N than larger populations; 2) there might be more variability among small populations in temporal variance N_b and/or N_b/N among small populations (*sensu* Wood et al. 2014).

We used coefficient of variations (CV: see Appendix A for sample calculations) as a standardized measure of relative variability (spatial and temporal) in N_b, N and N_b/N. To compare spatial and temporal dispersion in N_b, we examined boxplots of spatial and temporal CVs. To investigate whether CV in N_b and N_b/N was related to N, we plotted the temporal CV against harmonic mean of N for each stream. We then used simple linear regression to test whether there was a significant correlation between the average N and the temporal CV in N_b and N_b/N. To investigate whether there was more variability in temporal CVs among small populations, we used White's test (White 1980) implemented in the lmtest package (Zeileis & Hothorn 2002) to check for significant heteroscedasticity of the residuals in the fitted linear models.

Results

Intra-population genetic variation

CR brook trout populations were moderately genetically diverse with an average number of alleles per locus ranging from 2.08 to 7.18 alleles and an average heterozygosity from 0.15 to 0.60 (Table 1, Appendix A). For the 71 cohorts sampled, significant LD was detected in only 3.1% of 4686 LD tests (71 cohorts*66 pairwise locus comparisons) after Bonferroni correction and removing locus *Sco218*. Significant departures from HWE occurred in 4.6% of the 852 (71 cohorts*12 locus) tests after Bonferroni correction.

N, N_b and N_b/N ratios

Across the 12 populations, harmonic mean *N* ranged from 65-7801 with annual estimates ranging from 49 to 10032; harmonic mean N_b ranged from 5-355 with annual estimates ranging from 3-567 (table 1, Fig.2, Appendix A). On average, we had 5.25 *N* (range 4-6) and 5.25 N_b (range 4-7) estimates per population. We produced 43 N_b/N ratios by matching N_b estimates derived from YOY samples to *N* of the previous year; N_b estimates are derived from YOY shortly after emergence and therefore primarily reflect the parental generation (Waples 2005). Harmonic mean N_b/N ranged from 0.02-0.31 with annual estimates ranging from 0.01-0.45. On average, we had 3.6 N_b/N (2-5) ratios per population.

Effect of *N*, population, breeding year, and environmental variables on N_b

The LMMs showed a strong positive correlation between N_b and *N* and significant correlations between N_b and specific environmental variables. Indeed, the models describing the greatest amount of variability in N_b included *N* alone or with stream length and flow as explanatory variables (Table 2 and 3, Fig.3). As *N* and stream length increased, N_b increased as well; N_b also increased with flow, but most of the variability could be explained by stream size alone as shown by a reduced model only including stream size as an explanatory variable (Table 3). The population-specific term did not explain a large amount of variability in N_b in the models including N or stream length and flow as explanatory variables; however, the proportion of the variance explained increased two-fold in the model including only stream length as an explanatory variable (Table 3).

The relationship between N_b and stream length or N appeared linear on a log-log scale (Fig.3), suggesting that untransformed N_b and N or stream length follow a fractional power trend: as N and stream length increased, N_b increased at a decreasing rate. This provides some support for the hypothesis that N_b/N tend to be greater in small populations (genetic compensation).

Finally, LMM also suggested that there was a significant correlation between N and stream size (Appendix A), which further suggested that stream length could potentially be used to infer N or N_b. Indeed, the best fit LMM describing N only included stream length as an explanatory variable (Appendix A). Longer streams harboured larger brook trout populations, and up to 65% of the variability in N could be explained by stream length alone.

Relative trends in N, N_b and N_b/N within populations

Investigation of the temporal variation in N, Nb and Nb/N within population suggested a lack of temporal correlation between N and Nb, and provided support for a decrease in N_b/N following an increase in population size (genetic compensation). Linear models showed no relationship between a change in N and a change in N_b, or between a change in N and a change in N_b/N ($p=0.76$ and 0.32, respectively, df=21). However, this lack of relationship was mainly driven by a single, extreme outlier (a nearly 10-fold increase in N_b in a very small population, LC, in 2013 observable in Fig.2 as an increase in adult recruitment in 2015). Without the outlier, percent changes in N and N_b remained uncorrelated (Fig.4a; $p=0.12$, df=21), yet percent changes in N and N_b/N became significantly negatively correlated (Fig.4b; $p<0.001$, df=21, $r^2=0.47$). The relationship between changes in N and changes in N_b/N was not proportional: on average, N_b/N was expected to decrease by -1.55% per 1% increase in N.

Patterns of spatial and temporal variance in N_b, N and N_b/N

Across populations, spatial variability was significantly greater than temporal variability as shown by the non-overlapping notches of the boxplot. Spatial CVs in N, N_b and N_b/N were respectively 5.1-fold, 2.5-fold and 2.3-fold greater than temporal variation within populations (Fig.5a).

The linear model showed no relationship between N or N_b/N and temporal CVs (Fig.5b,c: $p=0.15$ and 0.77, df=10), suggesting that there was no correlation between population size and

variability in N_b and N_b/N. Finally, we had some evidence for increased variability among small populations in temporal variances; although the White's test was only significant for the temporal variability in N_b/N (Fig,5c; p=0.04,df=1), a similar trend was observed for temporal variability in N_b (Fig.5b; p=0.22,df=1).

Discussion

We set out to empirically evaluate the spatiotemporal relationship between N and N_b in relation to different environmental variables. To do so, we derived 63 N and N_b estimates from 12 pristine brook trout populations between 2008 and 2015. Our time series of data suggests that under certain biological and sampling conditions, conservation resources could be saved on the estimation of N or N_b to infer the other. We also provide some evidence for density-dependent constraints (genetic compensation) and environmental constraints on N_b (flow and stream length) and N (stream length).

Patterns of spatio-temporal variation in N_b and N

Our study revealed evidence of a strong correlation between N and N_b over a broad range of population sizes (N=65-7801 and N_b=5-355). Indeed, most of the variability in N_b could be accounted for by N or stream length alone. That is, larger habitats tended to harbour larger populations that were more genetically diverse. This result is consistent with island biogeography models at the species level (MacArthur & Wilson 1967) and other studies documenting correlations between habitat patch size, demographic and genetic population sizes within salmonid species (Whiteley *et al.* 2013; Wood *et al.* 2014) and small mammal species (Pardini *et al.* 2005). The strength and consistency of the relationships between N, N_b and stream length across populations and years in our study were especially intriguing given the variety of habitat conditions, trout population sizes, and trout life history diversity (e.g. adult body size, age- and size-at-maturation) among Cape Race streams (e.g. Hutchings 1993; Belmar-Lucero *et al.* 2012; Wood *et al.* 2014).

What ours and several other studies did not find was a correlation between N and N_b when they fluctuate temporally within a population (Araki *et al.* 2007; Ardren and Kapuscinski 2003; Duong *et al.* 2013; Johnstone *et al.* 2012; Palstra, O'Connell, and Ruzzante 2009; Serbezov *et al.* 2012; Whiteley *et al.* 2015). In fact, N_b appeared more temporally constrained than N. This discrepancy led to instability in the N_b/N ratio within populations; as N increased, N_b/N decreased, and vice-versa.

We also did not find any evidence for a consistent, directional difference between variability in N_b and/or N_b/N between small and large populations; however, small populations had more varying temporal variability in N_b/N ratios than large populations. Put another way, some small populations were highly variable in N_b/N while others remained relatively stable over time. Consequently, our results also strengthen insights from previous studies suggesting that small populations might be exposed to more variable selective pressures than larger populations (Wood *et al.* 2014). However, these trends were perhaps influenced by a reduced precision in the N_b estimates due to small sample sizes in some small or very small populations; small CR streams experience high variability in recruitment throughout the years, and despite intensive sampling it was sometimes extremely challenging to acquire large sample sizes from some of the streams.

Our results on the relationship between N_b and N therefore provide a mixed perspective on the utility of using N or N_b to infer the other. On one hand, because N_b is quite different in small and large populations and is also more spatially than temporally variable, N or N_b could be used to provide a reliable estimate for the other to assess relative population sizes. To a lesser extent, the strong relationship between N_b, N, a simple metric of habitat size (stream length) also suggests that stream size might serve as a useful proxy to predict N or N_b. This can be important, particularly in applications where the populations of interest are in remote or northern areas and hence not easily sampled. Nevertheless, the lack of strong temporal association between N_b and N and instability of N_b/N within populations suggest that inferring one variable from the other may be of limited use when the interest is to monitor population trends over time, perhaps especially for small populations of greater conservation concern because as our data suggests, their N_b - N dynamics may be less predictable than larger populations.

Finally, in many species, correctly linking N_b and N requires multiple sampling events, including over different years (Palstra and Fraser 2012). In many salmonids for example, and as illustrated in our study, juveniles and adults have to be sampled in t and t-1 years, respectively. In such situations, an important practical consideration is the extent to which the relationship between N_b and N might change when their estimates are incorrectly linked based on sampling conducted at the same time (e.g. in our case, summer). When we conducted an exploratory analysis on the relationship between unlinked N_b and N (sampling conducted in the same year), the relationship

remained essentially the same as in the correctly linked analyses (Appendix A). In general, we do not advocate the incorrect linking of N_b and N. Nevertheless, depending on the goal of a research study, these results suggest that resources (transportation and field costs, time) could be saved by reducing the number of sampling events, whilst obtaining reasonable information on demographic and genetic population size.

Given the aforementioned results, we propose four equations as a basis to infer relative genetic or demographic population sizes from the other or from stream length (L) in brook trout; 1) $\ln(N_b)=0.6760*\ln(N)–0.6256$; 2) $\ln(N)=1.4794*\ln(N_b)+0.9254$; 3) $\ln(N)=1.1903*\ln(L)-1.6228$; and 4) $\ln(N_b)=0.7309*\ln(\text{stream length})-1.2535$ (see Supplementary material for methodology and 95% confidence intervals). Extension of these conversions between N, N_b and stream length to other species with diverse life histories are currently under investigation.

Genetic compensation

The data collected on CR brook trout were consistent with the hypothesis that N_b/N ratios tend to be larger in small populations (genetic compensation). Among populations, N_b increased at a decreasing rate as N became larger; the two parameters were related linearly on a log-log scale, suggesting that the untransformed variables follow a fractional power trajectory. While our CR streams encompass a broad range of population and habitat sizes relative to one another, CR streams might be considered small in a broader context. Nevertheless, mean N and N_b included in this study do not significantly differ from other salmonid research, suggesting that our system is a good representation of the relationship between N and N_b in salmonid species. Of course, an analysis including much larger populations might be necessary to predict the shape of the relationship between N_b and N for species having typically large populations, such as marine fishes.

Further compelling evidence for the occurrence of genetic compensation came from the strong negative correlation between temporal changes in N and N_b/N within populations. Given the weak correlation between temporal changes in N and N_b, N_b must have remained relatively stable as N increased, thus reducing N_b/N. Most previous studies used correlations between the ratio

N_b/N and N as a means to investigate genetic compensation (see Jackson & Somers 1991 on spurious correlations); in avoiding this, our results strengthen the indications from previous works in providing, to our knowledge, a first concrete empirical example of genetic compensation in salmonids.

The presence of the significant correlation between stream length, N and N_b also provides a potential mechanism for the occurrence of genetic compensation. In salmonids, the initial mechanism proposed, based on competitive interference and density-dependence, is that an increase in population size increases the variance in reproductive success (Chebanov 1991). However, Whiteley et al. (2013, 2015) found no indication of density-dependent changes on indicators of reproductive success in brook trout. Additionally, we found no significant correlations between spawning habitat and stream size at CR (Appendix B), suggesting that spawning site availability may not limit the number of breeding adults.

An alternative mechanism for genetic compensation could be that density-dependent effects on individual growth result in temporal variation in important life history traits likely to affect N_b or N_b/N, such as fecundity or size- or age-at-maturity. At small population sizes or following a sharp decline in abundance, decreased adult density and increased food availability could result in higher growth, increased maturation rates and egg production (Rose *et al.* 2001), thus allowing small populations to maintain a relatively high N_b/N. In CR streams, we observed that in years where adult abundance was high, recruits tended to be smaller in size; we hypothesise that a certain proportion may have poorer spawning success due to poor body condition, or experience other physiological stresses that impact maturation or gamete production or quality. Partial support from this hypothesis comes from a recent study suggesting that only two life history variables could explain most of the variability in N_b/N across various taxa (Waples *et al.* 2013) and a simulation study within a salmonid system (Vincenzi *et al.* 2008).

Effect of the environment on N_b

In agreement with other recent studies showing that environmental variables can influence N_b (e.g. Whiteley et al. 2015), results from CR populations showed that flow had an effect on N_b , but only when stream size was used as a proxy for N; indeed, N_b tended to be higher when

summer flow was high. This suggests that N_b provides valuable insight into the reproductive biology of a species, and may therefore be important in its own right for understanding the eco-evolutionary dynamics of wild populations. However, it is noteworthy that larger streams tend to be characterized by faster flow on CR. Thus, the relationship between flow and N_b could be partially attributed to the relationship with stream length. More importantly, our data are also congruent with the idea that conservation actions restoring or maintaining habitat quantity and/or quality could be effective at improving the likelihood of wild population persistence. For instance, preventing further habitat degradation or fragmentation could help increase and/or maintain both N_b and N, thus improving long-term persistence potential.

Evolutionary and conservation implications

Empirical studies of the relationship between N and N_b among vertebrate populations spanning a large gradient of population sizes are rare. In a comprehensive time series including 12 brook trout populations, our results provide a mixed perspective on the utility of using N or N_b to infer the other. On one hand, there is a strong association between demographic and genetic population sizes, as well as between these parameters and a simple metric of habitat availability (stream length) when compared among populations; additionally, spatial variation among populations in N_b/N and N was greater than temporal variation within populations. Collectively, these findings suggest that the relative abundance of different populations can be inferred from N_b and habitat size, from a spatial comparison standpoint. For instance, non-invasive tissue/hair sampling could be used to estimate N_b and derive the relative N of remote (e.g. northern) populations, even without correctly linking N_b to N. Such knowledge could provide guidelines to set conservation targets. However, the lack of temporal correlation between N_b and N and the temporal instability of N_b/N ratios within populations suggests that monitoring of N_b or N from the other within populations is very challenging.

The observed differences in mean N_b/N between small and large populations and the increased variability in N_b/N variance among small populations also has potential ramifications in a global context where wild populations tend to become increasingly small and fragmented. These findings suggest that i) ecological interactions might consistently differ between small and large populations and that ii) small populations might be more heterogeneous than large populations

with respect to their selective pressures and likelihood of persistence in a dynamic and ever-changing world. Knowledge of which factors may favour small population persistence could therefore benefit conservation managers.

References

Anderson EC, Dunham KK (2008) The influence of family groups on inferences made with the program Structure. *Molecular ecology resources*, **8**, 1219–29.

Araki H, Waples RS, Ardren WR, Cooper B, Blouin MS (2007) Effective population size of steelhead trout: influence of variance in reproductive success, hatchery programs, and genetic compensation between life-history forms. *Molecular ecology*, **16**, 953–66.

Ardren WR, Kapuscinski AR (2003) Demographic and genetic estimates of effective population size (Ne) reveals genetic compensation in steelhead trout. *Molecular Ecology*, **12**, 35–49.

Belmar-Lucero S, Wood J, Scott S *et al.* (2012) Concurrent habitat and life history influences on effective/census population size ratios in stream-dwelling trout. *Ecology and evolution*, **2**, 562–73.

Berry O, Kirkwood R (2010) Measuring Recruitment in an Invasive Species to Determine Eradication Potential. *Journal of Wildlife Management*, **74**, 1661–1670.

Blanchfield PJ, Ridgway MS (1997) Reproductive timing and use of redd sites by lake-spawning brook trout (Salvelinus fontinalis). *Canadian Journal of Fisheries and Aquatic Sciences*, **54**, 747–756.

Blanchfield PJ, Ridgway MS (2005) The relative influence of breeding competition and habitat quality on female reproductive success in lacustrine brook trout (Salvelinus fontinalis). *Canadian Journal of Fisheries and Aquatic Sciences*, **62**, 2694–2705.

Charlier J, Laikre L, Ryman N (2012) Genetic monitoring reveals temporal stability over 30 years in a small, lake-resident brown trout population. *Heredity*, **109**, 246–53.

Chebanov N (1991) Effect of spawner density on spawning success, egg survival and size structure on the progeny of the sockeye salmon, Oncorhynchus nera. *Journal of Ichtyologie*, **31**, 103:109.

Curry RA, Neakes DLG (1995) Groundwater and the selection of spawning sites by brook trout (Salvelinus fontinalis). *Canadian Journal of Aquatic Sciences*, **52**, 1733–1740.

Danzmann RG, Ii RPM, Jones MWJ, Bernatchez L, Ihssen PE (1998) A major sextet of mitochondrial DNA phylogenetic assemblages extant in eastern North American brook trout

(Salvelinus fontinalis): distribution and postglacial dispersal patterns. *Canadian Journal of Zoology*, **76**, 1300–1318.

Dowling TE, Turner TF, Carson EW *et al.* (2014) Time-series analysis reveals genetic responses to intensive management of razorback sucker (Xyrauchen texanus). *Evolutionary applications*, **7**, 339–54.

Duong TY, Scribner KT, Forsythe PS, Crossman J a, Baker E a (2013) Interannual variation in effective number of breeders and estimation of effective population size in long-lived iteroparous lake sturgeon (Acipenser fulvescens). *Molecular ecology*, **22**, 1282–94.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular ecology*, **14**, 2611–20.

Fraser DJ, Bernatchez L (2005) Adaptive migratory divergence among sympatric brook char populations. *Evolution*, **59**, 611–624.

Fraser DJ, Debes P V, Bernatchez L, Hutchings JA, Fraser DJ (2014) Population size , habitat fragmentation , and the nature of adaptive variation in a stream fish. *Proceedings of the Royal Society of London B: Biological Sciences*, **281**, 20140370.

Fraser DJ, Jones MW, McParland TL, Hutchings J a. (2007) Loss of historical immigration and the unsuccessful rehabilitation of extirpated salmon populations. *Conservation Genetics*, **8**, 527–546.

Guillementte F, Vallee C, Bertolo A, Magnan P (2011) The evolution of redd site selection in brook charr in different environments : Same cue , same benefit for fitness. *Freshwater Biology*, **56**, 1017–1029.

Hakala JP, Hartman KJ (2004) Drought effect on stream morphology and brook trout (Salvelinus fontinalis) populations in forested headwater streams . *Hydrobiologia*, **515**, 202–213.

Hua YG, Orban L (2005) A simple and affordable method for high throughput DNA extraction from animal tissues for PCR. *Electrophoresis*, **26**, 3081–3083.

Hutchings JA (1993) Adaptive Life Histories Effected by Age-Specific Survival and Growth Rate. *Ecology*, **74**, 673–684.

Jackson DA, Somers KM (1991) The spectre of "spurious" correlations. *Oecologia*, **86**, 147–151.

Johnstone DL, O'Connell MF, Palstra FP, Ruzzante DE (2012) Mature male parr contribution to the effective size of an anadromous Atlantic salmon (Salmo salar) population over 30 years. *Molecular Ecology*, **22**, 2394–2407.

Kanno Y, Letcher BH, Hitt NP *et al.* (2015) Seasonal weather patterns drive population vital rates and persistence in a stream fish. *Global change biology*, **21**, 1856–70.

Kenward MG, Roger JH (1997) Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, **53**, 983–97.

Lenth RV, Herva M (2015) lsmeans: Least-Squares Means. R package version 2.17. *http://CRAN.R-project.org/package=lsmeans*.

Letcher BH, Schueller P, Bassar RD *et al.* (2014) Robust estimates of environmental effects on population vital rates: an integrated capture-recapture model of seasonal brook trout growth, survival and movement in a stream network. *The Journal of animal ecology*, **84**, 337–352.

MacArthur R, Wilson E (1967) *The theory of island biogeography.* Princeton University Press.

Nakagawa S, Schielzeth H (2013) A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, **4**, 133–142.

Osborne MJ, Davenport SR, Hoagstrom CW, Turner TF (2010) Genetic effective size, N(e), tracks density in a small freshwater cyprinid, Pecos bluntnose shiner (Notropis simus pecosensis). *Molecular ecology*, **19**, 2832–44.

Palstra FP, Fraser DJ (2012) Effective/census population size ratio estimation: a compendium and appraisal. *Ecology and evolution*, **2**, 2357–65.

Palstra FP, O'Connell MF, Ruzzante DE (2009) Age Structure, Changing Demography and Effective Population Size in Atlantic Salmon (Salmo salar). *Genetics*, **182**, 1233–1249.

Palstra FP, Ruzzante DE (2008) Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? *Molecular Ecology*, **17**, 3428–3447.

Pankhurst NAW, Munday PBL (2011) Effects of climate change on fish reproduction and early life history stages. *Marine and Freshwater Research*, **62**, 1015–1026.

Pardini R, De Souza SM, Braga-Neto R, Metzger JP (2005) The role of forest structure, fragment size and corridors in maintaining small mammal abundance and diversity in an Atlantic forest landscape. *Biological Conservation*, **124**, 253–266.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, **28**, 2537–2539.

Perrier C, April J, Cote G, Bernatchez L, Dionne M (2015) Effective number of breeders in relation to census size as management tools for Atlantic salmon conservation in a context of stocked populations. *Conservation Genetics*, **17**, 31–44.

Petersen C (1985) The Yearly Immigration of Young Plaice Into the Limfjord From the German Sea. *Report of the Danish Biological Station*, **5**, 5–84.

Pinheiro J, Bates D, DebRoy S, Sarkar D, Team the RDC (2013) nlme: linear and nonlinear mixed effects models. R pac. *R package version 3.1-113*.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics Society of America*, **155**, 945–959.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Rose KA, Cowan JH, Winemiller KO, Myers R a., Hilborn R (2001) Compensatory density dependence in fish populations: Importance, controversy, understanding and prognosis. *Fish and Fisheries*, **2**, 293–327.

Schwartz MK, Luikart G, Waples RS (2007) Genetic monitoring as a promising tool for conservation and management. *Trends in ecology & evolution*, **22**, 25–33.

Serbezov D, Jorde PE, Bernatchez L, Olsen EM (2012) Life history and demographic determinants of effective / census size ratios as exemplified by brown trout (Salmo trutta). *Evolutionary Applications*, **5**, 607–618.

Tallmon D a, Gregovich D, Waples RS *et al.* (2010) When are genetic methods useful for

estimating contemporary abundance and detecting population trends? *Molecular ecology resources*, **10**, 684–92.

Team RC (2013) R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria.*

Vincenzi S, Crivelli AJ, Jesensek D, De Leo GA (2008) The role of density-dependent individual growth in the persistence of freshwater salmonid populations. *Oecologia*, **156**, 523–534.

Wang J (2004) Sibship Reconstruction From Genetic Data With Typing Errors. *Genetics Society of America*, **1979**, 1963–1979.

Waples RS (2002) Effective size of fluctuating salmon populations. *Genetics*, **161**, 783–91.

Waples R (2005) Genetic estimates of contemporary effective population size : to what time periods do the estimates apply ? *Molecular Ecology*, **14**, 3335–3352.

Waples RS, Antao T, Luikart G (2014) Effects of Overlapping Generations on Linkage Disequilibrium Estimates of Effective Population Size. *Genetics*, **197**, 769–780.

Waples RS, Do C (2010) Linkage disequilibrium estimates of contemporary N e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, **3**, 244–262.

Waples RS, Luikart G, Faulkner JR, Tallmon DA (2013) Simple life-history traits explain key effective population size ratios across diverse taxa. *Proceedings of the Royal Society of London B: Biological Science*, **280**, 20131339.

Warren DR, Robinson JM, Josephson DC, Sheldon DR, Kraft CE (2012) Elevated summer temperatures delay spawning and reduce redd construction for resident brook trout (Salvelinus fontinalis). *Global Change Biology*, **18**, 1804–1811.

White H (1980) A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, **48**, 817–838.

Whiteley AR, Coombs JA, Cembrola M *et al.* (2015) Effective number of breeders provides a link between interannual variation in stream flow and individual reproductive contribution in a stream salmonid. *Molecular Ecology*, **24**, 3585–3602.

Whiteley AR, Coombs JA, Hudy M *et al.* (2013) Fragmentation and patch size shape genetic structure of brook trout populations. *Canadian Journal of Fisheries and Aquatic Sciences*, **688**, 678–688.

Wood JLA, Belmar-Lucero S, Hutchings IJ a, Fraser DJ (2014) Relationship of habitat variability to population size in a stream fish. *Ecological applications : a publication of the Ecological Society of America*, **24**, 1085–100.

Wood JLA, Tezel D, Joyal D, Fraser DJ (2015) Population size is weakly related to quantitative genetic variation and trait differentiation in a stream fish. *Evolution*, **69**, 2303–2318.

Xu C, Letcher BH, Nislow KH (2010) Context-specific influence of water temperature on brook trout growth rates in the field. *Freshwater Biology*, **55**, 2253–2264.

Zeileis A, Hothorn T (2002) Diagnostic Checking in Regression Relationships. *R News*, **2**, 7–10.

Tables

Table 1: Population summary statistics from twelve brook trout populations from Cape Race (Newfoundland, Canada) sampled between 2008-2015 (see details per year/cohort in Appendix A).

	#*N*	#N_b	Mean *N*	Mean N_b	Mean S	Total S	A	Ae	Ho	He
UC	6	5	65 (49 - 79)	20 (13-48)	19 (5-35)	115	3.09 (2.42-4.33)	1.90 (1.57-2.42)	0.39 (0.28-0.54)	0.41 (0.31-0.53)
HM	4	4	66 (52 - 80)	5 (3-5)	25 (4-61)	126	2.08 (1.42-3.17)	1.39 (1.09-1.66)	0.15 (0.06-0.20)	0.22 (0.06-0.52)
DY	6	5	116 (84 - 179)	10 (3-34)	26 (9-35)	153	3.85 (3.00-4.33)	2.47 (2.06-2.72)	0.54 (0.49-0.61)	0.52 (0.45-0.57)
LC	6	6	338 (250 - 798)	31 (11-117)	43 (11-60)	299	3.82 (2.25-5.58)	2.23 (1.70-2.96)	0.44 (0.40-0.50)	0.48 (0.37-0.59)
LO	5	5	470 (372 - 625)	44 (23 - 188)	31 (25-39)	155	4.35 (3.75-4.67)	2.62 (2.34-2.81)	0.56(0.47-0.62)	0.56 (0.51-0.60)
WC	6	6	783 (530 - 1148)	31 (21-52)	98 (48-153)	586	5.04 (3.83-6.17)	2.71 (2.21-3.12)	0.52 (0.46-0.56)	0.55 (0.48-0.58)
STBC	6	5	917 (587 - 1405)	28 (14-54)	63 (40-97)	444	2.48 (1.83-3.50)	1.54 (1.48-1.61)	0.27(0.24-0.30)	0.28 (0.25-0.31)
LB	4	5	1184 (877 - 1383)	52 (34-83)	55 (52-58)	274	5.30 (4.75-5.75)	2.94 (2.74-3.08)	0.54 (0.50-0.60)	0.56 (0.53-0.57)
CC	4	5	1862 (1471 - 5246)	74 (65-99)	104 (71-163)	520	3.81 (3.33-4.58)	2.09 (2.03-2.24)	0.43 (0.40-0.47)	0.42 (0.40-0.46)
UO	6	6	2569 (1949 - 3835)	62 (41-87)	113 (36-237)	679	5.92 (5.25-6.50)	3.25 (3.04-3.32)	0.60 (0.56-0.64)	0.62 (0.61-0.63)
BC	6	4	4693 (4044 - 6132)	355 (267-567)	123 (62-223)	738	4.60 (4.17-5.25)	2.23 (2.17-2.29)	0.46 (0.44-0.49)	0.45 (0.44-0.46)
WN	4	7	7801 (6713 - 10032)	178 (110-267)	101 (56-160)	707	7.18 (6.50-8.00)	3.29 (3.05-3.53)	0.57 (0.56-0.60)	0.59 (0.56-0.60)

Number of annual *N* and N_b estimates per population (#*N* and #N_b), harmonic means of *N* and N_b (mean *N* and Mean N_b), mean annual sample size (Mean S), total number of genotyped YOY (Total S), mean annual number of alleles per locus (A), mean annual effective number of alleles per locus (Ae), mean annual observed heterozygosity (Ho), mean annual expected heterozygosity (He). Ranges across years are between parentheses.

Table 2: Test statistics for the final model terms predicting N_b in 12 brook trout populations from Cape Race, Newfoundland, Canada.

	N model			Stream length		
Parameter	Df	F value	p value	df	F value	p value
Ln (*N*) or Ln (stream size)	1,14.0	20.7	<0.001	1,11.3	7.72	0.017
Velocity	-	-	-	1,20.6	7.21	0.014

Table 3: Mixed-effect modeling of the effect of *N* and environmental variables on N_b among 12 brook trout populations from Cape Race, Newfoundland, Canada.

	N model	Full stream length	Restricted stream length
Fixed effects estimates (SE)			
Intercept	3.82 (0.15)	3.83 (0.18)	3.84 (0.21)
N or stream length	0.85 (0.15)	0.55 (0.20)	0.79 (0.21)
Flow	-	0.54 (0.19)	-
Random effects variance components (Sd)			
Population	0.17 (0.41)	0.22 (0.47)	0.48 (0.69)
residual	0.42 (0.65)	0.41 (0.64)	0.32 (0.56)
R^2_m	0.54	0.53	0.44
R^2_c	0.67	0.69	0.77
n	43	32	63

SE= standard error; Sd = standard deviation; R^2_m = marginal R square (fixed effects only); R^2_c = conditional R square (fixed and random effects); n=number of observations

Figures

Fig. 1: Cape Race streams. From West to East; 1) Lower Coquita (LC), 2) Upper Coquita (UC), 3) Hermitage (HM), 4) Bob's Cove (BC), 5) Still There By Chance (STBC), 6) Whale Cove (WC), 7) Ditchy (DY), 8) Lower O'Beck (LO), 9) Upper O'Beck (UO), 10) Watern (WN), 11) Blackfly (LB), 12) Cripple Cove (CC).

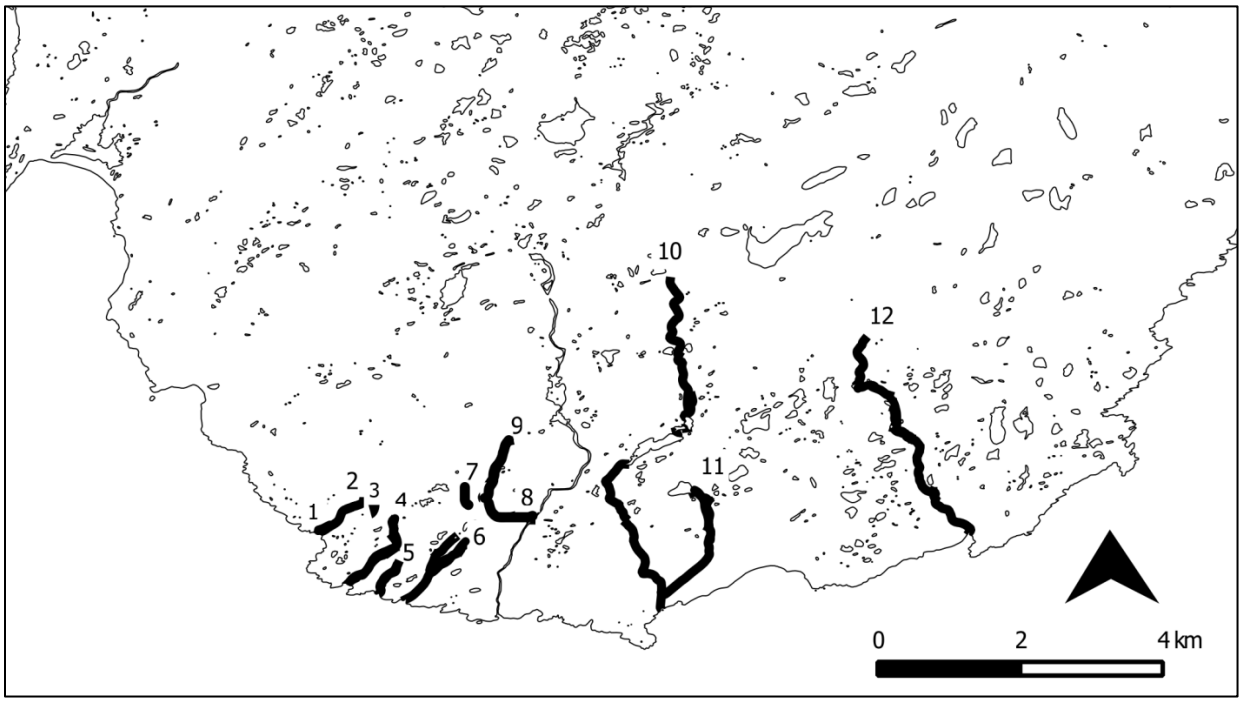


Fig.2: Temporal trends in N and N_b estimates within 12 brook trout populations from Cape Race, Newfoundland, Canada. N was estimated using mark-recapture methods and N_b was derived from LDNe. Upper CI including infinity were not represented.

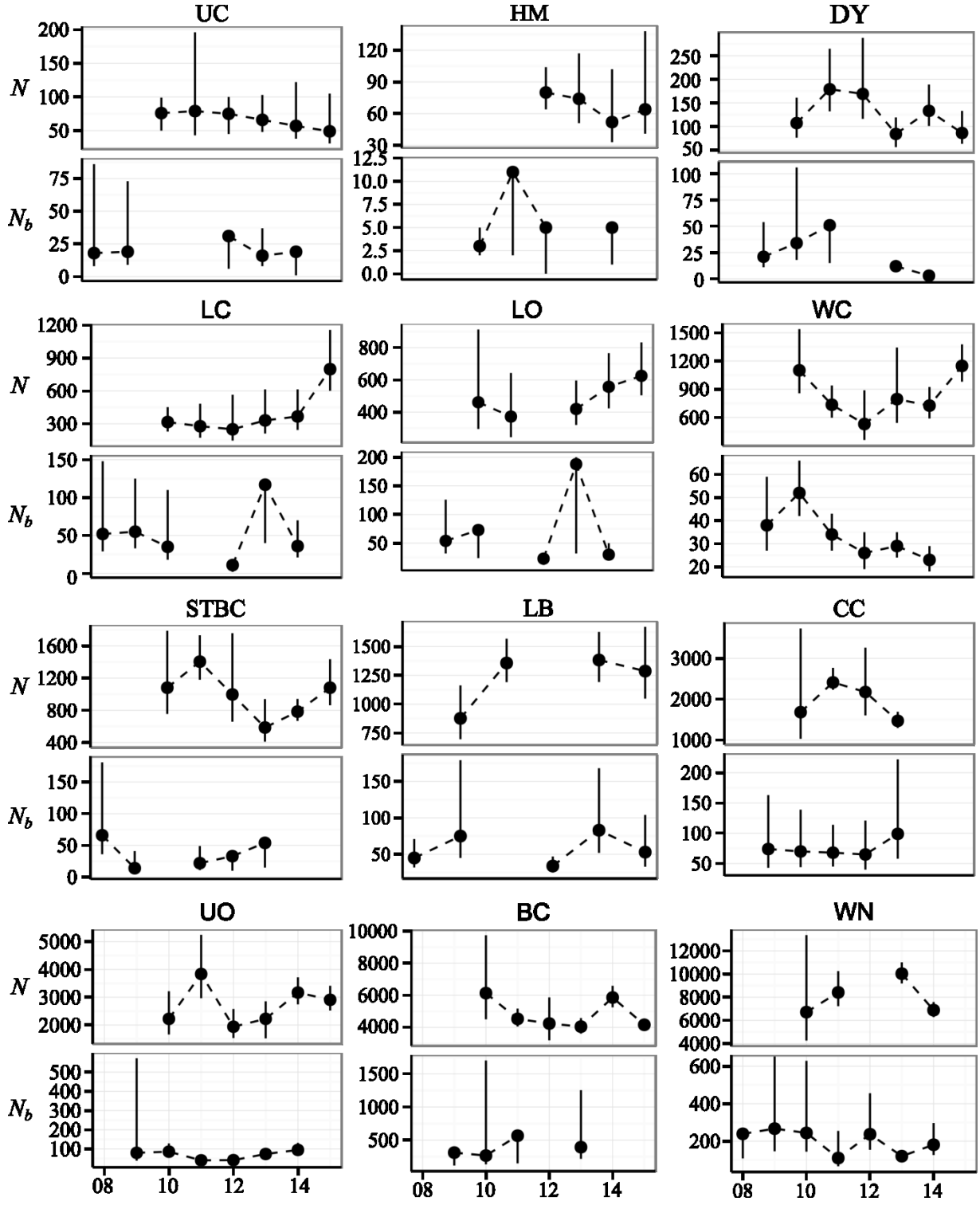


Fig.3: Relationship of N_b with N and stream length in 12 brook trout populations from Cape Race, Newfoundland, Canada.

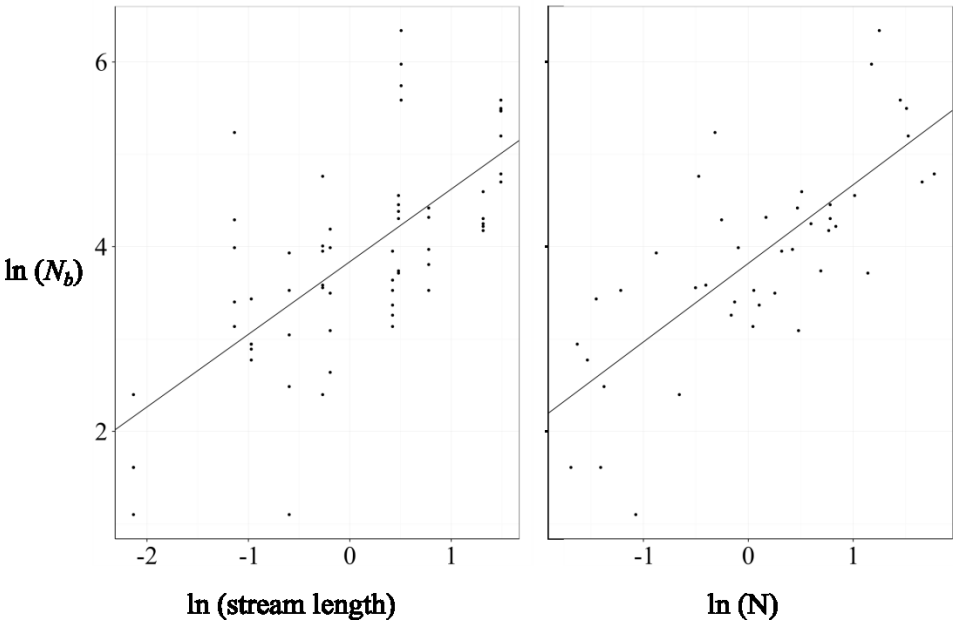


Fig.4: Relationships between percent changes in annual N and the percent change in annual N_b or the percent change in annual N_b/N after removal of a far outlier.

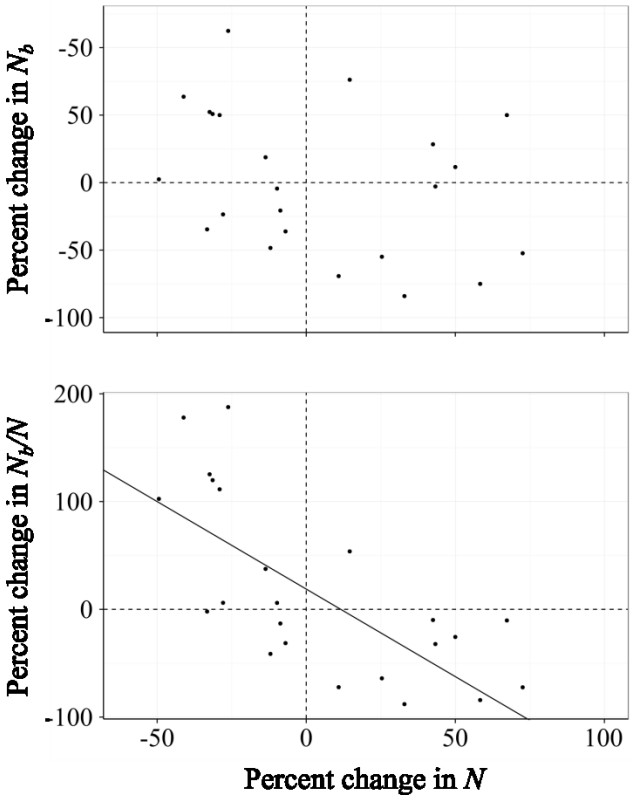
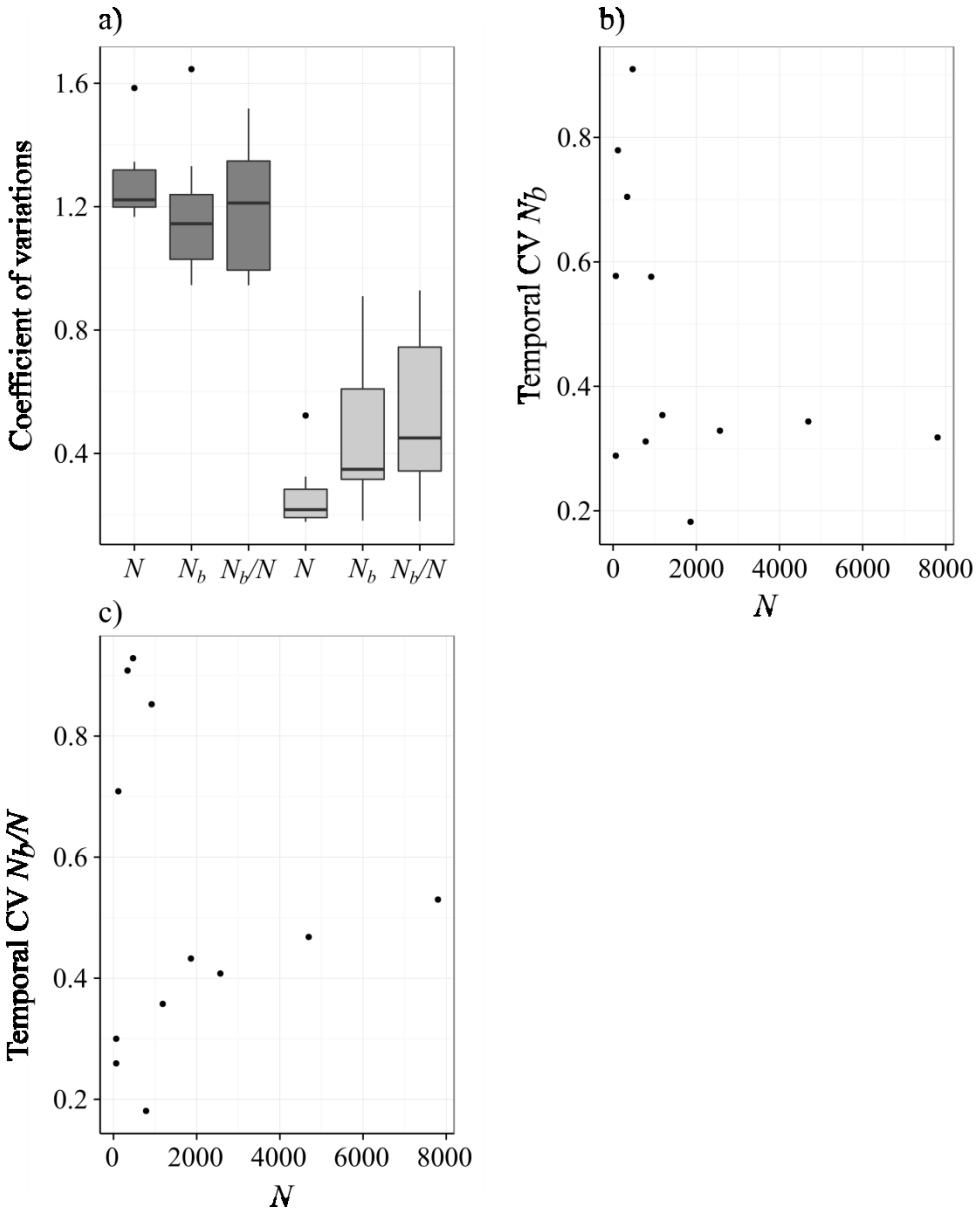


Fig.5: **a)** Among population (spatial) variation (dark grey) and within population (temporal) variation (light grey) in N, N_b, and N_b/N; **b)** relationship between population size and variance in N_b; **c)** relationship between harmonic mean census population size and variance in N_b/N. Coefficient of variations (CV) were used as a standardized metric of variability.



Appendix A

Table A1: Annual population summary statistics for 12 Cape Race, Newfoundland, Canada, brook trout populations sampled between 2008-2015.

Table A2: Environmental data summary statistics for 12 CR brook trout populations between 2010-2014.

Table A3: STRUCTURE analysis of population subdivision in four brook trout drainages from Cape Race, Newfoundland

Table A4: Test statistics for the final model terms predicting N_b in 12 brook trout populations from Cape Race, Newfoundland, Canada.

Table A5: Mixed-effect modeling of the relationship between N and stream length among 12 brook trout populations from Cape Race, Newfoundland, Canada.

Table A6: Temporal variability in N, N_b and N_b/N within 12 brook trout populations from Cape Race, Newfoundland.

Table A7: Spatial variability in N, N_b and N_b/N among 12 brook trout populations from Cape Race, Newfoundland.

Table A8: Mixed-effect modeling of the relationship between incorrectly linked N and N_b, sampled from the same season (based on adult tagging for N estimation and juvenile YOY sampling for N_b estimation)

Fig.A1: Relationship between N and stream size from a mixed model.

Fig.A2: Relationship between incorrectly linked N and N_b from a mixed model, sampled from the same season (based on adult tagging for N estimation and juvenile YOY sampling for N_b estimation).

Table A1: Annual population summary statistics for 12 brook trout populations from Cape Race, Newfoundland, Canada, sampled between 2008-2015.

Pop	Year	*N* (95% CI)	N_b (95% CI)	S	A	Ae	Ho	He
BC	2008	NA	∞ (166-∞)	62	4.25 (0.58)	2.29 (0.32)	0.48 (0.08)	0.46 (0.08)
BC	2009	NA	312 (117-∞)	95	4.50 (0.76)	2.29 (0.31)	0.49 (0.08)	0.46 (0.08)
BC	2010	6132 (4500-9739)	267 (135-1702)	105	5.25 (0.79)	2.21 (0.30)	0.47 (0.08)	0.45 (0.08)
BC	2011	4527 (4052-5167)	567 (150-∞)	151	4.17 (0.58)	2.17 (0.27)	0.44 (0.08)	0.45 (0.08)
BC	2012	4238 (3191-5867)	∞ (576-∞)	102	4.25 (0.62)	2.22 (0.30)	0.45 (0.08)	0.45 (0.08)
BC	2013	4044 (3610-4579)	394 (218-1254)	223	5.17 (0.73)	2.17 (0.30)	0.45 (0.08)	0.44 (0.08)
BC	2014	5859 (5246-6590)	NA	NA	NA	NA	NA	NA
BC	2015	4158 (3826-4546)	NA	NA	NA	NA	NA	NA
CC	2009	NA	74 (43-163)	76	4.00 (0.56)	2.07 (0.30)	0.43 (0.07)	0.42 (0.07)
CC	2010	1683 (1032-3731)	70 (44-139)	71	4.58 (0.57)	2.24 (0.31)	0.47 (0.07)	0.46 (0.07)
CC	2011	2412 (2231-2768)	68 (45-114)	163	3.33 (0.43)	2.07 (0.27)	0.41 (0.07)	0.42 (0.07)
CC	2012	2176 (1603-3261)	65 (40-121)	94	3.33 (0.47)	2.03 (0.26)	0.43(0.07)	0.42 (0.07)
CC	2013	1471 (1295-1693)	99 (58-222)	116	3.83 (0.37)	2.05 (0.30)	0.40 (0.07)	0.40 (0.07)
DY	2009	NA	21 (11-54)	26	4.25 (0.46)	2.72 (0.36)	0.61 (0.06)	0.57 (0.05)
DY	2010	107 (76-161)	34 (18-106)	35	4.33 (0.54)	2.50 (0.38)	0.53 (0.05)	0.53 (0.05)
DY	2011	179 (132-265)	51 (15-∞)	34	3.50 (0.42)	2.33 (0.27)	0.52 (0.05)	0.52 (0.05)
DY	2012	169 (116-288)	∞ (8.6-∞)	9	3.00 (0.30)	2.06 (0.25)	0.49 (0.06)	0.45 (0.05)
DY	2013	84 (56-119)	12 (9-18)	35	4.25 (0.68)	2.72 (0.56)	0.55 (0.05)	0.54 (0.05)
DY	2014	133 (101-189)	3 (2-7)	14	3.75 (0.43)	2.50 (0.39)	0.56 (0.06)	0.52 (0.05)
DY	2015	86 (63-133)	NA	NA	NA	NA	NA	NA
HM	2010		3 (2-5)	61	3.17 (0.62)	1.65 (0.25)	0.20 (0.06)	0.52 (0.05)
HM	2011	NA	11 (2-∞)	29	2.58 (0.72)	1.39 (0.18)	0.16 (0.07)	0.18 (0.07)
HM	2012	80 (64-104)	5 (0-∞)	17	1.42 (0.19)	1.09 (0.07)	0.06 (0.04)	0.06 (0.04)
HM	2013	74 (51-117)	∞ (0-∞)	4	1.67 (0.26)	1.37 (0.16)	0.20 (0.09)	0.18 (0.07)
HM	2014	52 (33-102)	5 (1-∞)	15	1.58 (0.31)	1.40 (0.21)	0.14 (0.07)	0.15 (0.08)
HM	2015	64 (41-138)	NA	NA	NA	NA	NA	NA
LB	2009	NA	45 (32-71)	54	5.33 (0.80)	2.91 (0.42)	0.54 (0.09)	0.55 (0.08)
LB	2010	877 (696-1163)	75 (45-179)	52	5.50 (0.76)	2.99 (0.40)	0.56 (0.08)	0.57 (0.07)
LB	2011	1357 (1191-1568)	NA	NA	NA	NA	NA	NA
LB	2012	NA	34 (25-47)	56	5.75 (0.82)	3.08 (0.46)	0.52 (0.07)	0.57 (0.07)
LB	2013	1383 (1192-1627)	83 (52-168)	58	5.17 (0.73)	2.97 (0.38)	0.60 (0.08)	0.56 (0.08)
LB	2014	1287 (1048-1671)	53 (33-104)	54	4.75 (0.62)	2.74 (0.34)	0.50 (0.08)	0.53 (0.08)
LC	2008	NA	52 (29-148)	48	5.58 (0.88)	2.96 (0.39)	0.50 (0.05)	0.59 (0.05)
LC	2009	NA	55 (33-125)	59	4.42 (0.88)	2.37 (0.22)	0.41 (0.05)	0.53 (0.05)
LC	2010	316 (229-452)	35 (18-110)	42	4.92 (0.82)	2.59 (0.41)	0.45 (0.07)	0.51 (0.07)
LC	2011	278 (173-483)	∞ (26-∞)	22	3.08 (0.31)	1.98 (0.21)	0.40 (0.07)	0.44 (0.05)

LC	2012	250 (147-565)	11 (2-∞)	11	2.25 (0.18)	1.70 (0.12)	0.48 (0.07)	0.37 (0.05)
LC	2013	331 (211-614)	117 (40-∞)	57	3.42 (0.48)	1.99 (0.21)	0.43 (0.06)	0.45 (0.05)
LC	2014	367 (244-614)	36 (21-70)	60	3.08 (0.34)	2.01 (0.17)	0.42 (0.06)	0.46 (0.04)
LC	2015	798 (601-1156)	NA	NA	NA	NA	NA	NA
LO	2009	NA	54 (32-126)	39	4.67 (0.69)	2.71 (0.37)	0.58 (0.05)	0.57 (0.05)
LO	2010	461 (295-913)	73 (24-∞)	25	4.67 (0.68)	2.81 (0.35)	0.62 (0.04)	0.60 (0.04)
LO	2011	372 (244-643)	NA	NA	NA	NA	NA	NA
LO	2012	NA	23 (12-62)	26	3.75 (0.63)	2.34 (0.35)	0.47 (0.08)	0.51 (0.06)
LO	2013	419 (320-596)	188 (32-∞)	30	4.17 (0.53)	2.48 (0.32)	0.54 (0.05)	0.54 (0.05)
LO	2014	557 (423-766)	30 (20-50)	35	4.50 (0.72)	2.76 (0.39)	0.57 (0.05)	0.57 (0.05)
LO	2015	625 (504-833)	NA	NA	NA	NA	NA	NA
STBC	2008	NA	66 (36-181)	93	3.50 (0.42)	1.61 (0.17)	0.26 (0.06)	0.31 (0.07)
STBC	2009	NA	14 (6-41)	42	3.08 (0.26)	1.59 (0.14)	0.30 (0.08)	0.31 (0.06)
STBC	2010	1081 (755-1788)	∞ (31-∞)	40	2.50 (0.29)	1.58 (0.16)	0.27 (0.07)	0.29 (0.07)
STBC	2011	1405 (1180-1732)	22 (11-49)	97	2.67 (0.38)	1.52 (0.15)	0.24 (0.06)	0.27 (0.07)
STBC	2012	997 (659-1758)	33 (10-401)	63	1.83 (0.27)	1.51 (0.16)	0.26 (0.07)	0.25 (0.07)
STBC	2013	587 (411-940)	54 (15-∞)	64	1.92 (0.29)	1.52 (0.17)	0.25 (0.07)	0.25 (0.07)
STBC	2014	784 (669-942)	∞ (37-∞)	45	1.83 (0.27)	1.48 (0.15)	0.28 (0.08)	0.25 (0.07)
STBC	2015	1081 (863-1435)	NA	NA	NA	NA	NA	NA
UC	2008	NA	18 (8-86)	19	4.33 (0.71)	2.42 (0.25)	0.54 (0.06)	0.53 (0.06)
UC	2009	NA	19 (9-73)	25	3.75 (0.49)	2.15 (0.23)	0.44 (0.07)	0.49 (0.04)
UC	2010	76 (50-99)	NA	NA	NA	NA	NA	NA
UC	2011	79 (43-196)	∞ (7-∞)	15	2.42 (0.29)	1.57 (0.15)	0.28 (0.07)	0.31 (0.06)
UC	2012	75 (45-100)	31 (6-∞)	16	2.67 (0.33)	1.71(0.22)	0.34 (0.08)	0.32 (0.07)
UC	2013	66 (48-103)	16 (8-37)	35	2.92 (0.34)	1.84 (0.19)	0.33 (0.07)	0.39 (0.07)
UC	2014	57 (38-122)	19 (1-∞)	5	2.42 (0.23)	1.72 (0.10)	0.40 (0.07)	0.39 (0.04)
UC	2015	49 (31-105)	NA	NA	NA	NA	NA	NA
UO	2009	NA	80 (39-572)	36	5.25 (0.75)	3.04 (0.41)	0.64 (0.05)	0.61 (0.05)
UO	2010	2222 (1660-3215)	86 (63-129)	93	6.17 (0.90)	3.32 (0.52)	0.62 (0.05)	0.63 (0.05)
UO	2011	3835 (2964-5245)	41 (33-52)	132	5.75 (0.80)	3.29 (0.55)	0.58 (0.05)	0.61 (0.05)
UO	2012	1940 (1532-2578)	42 (33-57)	76	5.83 (0.88)	3.20 (0.57)	0.56 (0.04)	0.62 (0.05)
UO	2013	2223 (1519-2853)	74 (63-88)	237	6.50 (1.07)	3.31 (0.55)	0.60 (0.06)	0.62 (0.05)
UO	2014	3168 (2743-3716)	95 (72-132)	105	6.00 (0.83)	3.32 (0.55)	0.59 (0.05)	0.62 (0.05)
UO	2015	2907 (2521-3412)	NA	NA	NA	NA	NA	NA
WC	2009	NA	38 (27-59)	48	5.00 (0.52)	2.80 (0.43)	0.54 (0.07)	0.56 (0.05)
WC	2010	1101 (857-1539)	52 (42-66)	108	6.17 (0.82)	3.12 (0.57)	0.56 (0.06)	0.58 (0.06)
WC	2011	735 (597-940)	34 (27-43)	153	5.25 (0.64)	2.63 (0.43)	0.50 (0.06)	0.53 (0.06)
WC	2012	530 (361-889)	26 (19-35)	64	3.83 (0.55)	2.21 (0.33)	0.46 (0.07)	0.48 (0.06)
WC	2013	795 (542-1343)	29 (24-35)	143	5.50 (0.66)	2.78 (0.39)	0.53 (0.06)	0.56 (0.06)
WC	2014	726 (589-924)	23 (18-29)	70	4.50 (0.52)	2.72 (0.37)	0.53 (0.06)	0.55 (0.06)
WC	2015	1148 (980-1377)	NA	NA	NA	NA	NA	NA

WN	2008	NA	239 (108-∞)	59	7.33 (1.18)	3.49 (0.58)	0.60 (0.08)	0.60 (0.07)
WN	2009	NA	267 (146-1016)	96	7.08 (1.06)	3.19 (0.44)	0.57 (0.08)	0.59 (0.08)
WN	2010	6713 (4256-13367)	244 (144-629)	100	7.17 (0.86)	3.37 (0.48)	0.57 (0.08)	0.58 (0.08)
WN	2011	8416 (7225-10255)	110 (66-255)	56	6.50 (0.97)	3.28 (0.49)	0.56 (0.08)	0.58 (0.08)
WN	2012	NA	237 (154-456)	119	6.67 (1.00)	3.05 (0.41)	0.57 (0.08)	0.56 (0.08)
WN	2013	10032 (9188-11016)	120 (97-150)	160	8.00 (1.16)	3.53 (0.56)	0.57 (0.08)	0.60 (0.08)
WN	2014	6883 (6283-7588)	181 (126-297)	117	7.50 (0.96)	3.11 (0.45)	0.57 (0.07)	0.59 (0.07)
WN	2015	NA	NA	NA	NA	NA	NA	NA

Adult census population sizes derived from mark-recapture methods (N), effective number of breeders (N_b) estimates derived from LDNe, sample size (S), number of alleles (A), effective number of alleles (Ae) defined as the number of equal frequency alleles it takes to achieve the expected heterozygosity (He), observed heterozygosity (Ho), and expected heterozygosity (He). Upper and lower 95% confidence intervals around N and N_b are between brackets: A, Ae, Ho and He represent the mean over 12 locus with the associated standard errors between parentheses.

Table A2: Environmental data summary statistics for 12 Cape Race brook trout populations between 2010-2014.

Population	Year	Stream length (m)	flow (Sd)	Temperature (Sd)
BC	2010	2050	0.22 (0.23)	13.01 (2.56)
BC	2011	2050	0.23 (0.22)	NA
BC	2012	2050	0.25 (0.15)	15.16 (2.40)
BC	2013	2050	0.22 (0.17)	13.09 (2.60)
BC	2014	2050	0.3 (0.22)	15.02 (2.60)
CC	2010	6318	0.23 (0.10)	14.98 (2.60)
CC	2011	6318	0.18 (0.14)	NA
CC	2012	6318	0.17 (0.09)	16.08 (2.35)
CC	2013	6318	0.15 (0.13)	14.6 (2.55)
CC	2014	6318	0.37 (0.18)	14.68 (2.64)
DY	2010	441	0.01 (0.02)	13.98 (2.10)
DY	2011	441	0.01 (0.02)	NA
DY	2012	441	0.08 (0.27)	15.58 (1.98)
DY	2013	441	0.07 (0.09)	14.5 (2.06)
DY	2014	441	0.01 (0.03)	14.97 (2.35)
HM	2012	52	0 (0)	15.72 (2.45)
HM	2013	52	0 (0)	NA
HM	2014	52	0 (0)	17.35 (2.24)
LB	2010	2993	0.16 (0.11)	14.08 (1.94)
LB	2011	2993	0.15 (0.14)	NA
LB	2013	2993	0.12 (0.14)	15.45 (1.33)
LB	2014	2993	0.16 (0.13)	11.91 (1.86)
LC	2010	698	0.12 (0.12)	11.11 (2.08)
LC	2011	698	0.15 (0.13)	NA
LC	2012	698	0.17 (0.34)	10.97 (1.59)
LC	2013	698	0.22 (0.23)	14.74 (1.48)
LC	2014	698	0.18 (0.15)	11.37 (1.55)

LO	2010	208	0.3 (0.19)	NA
LO	2011	208	0.24 (0.20)	NA
LO	2012	208	NA	NA
LO	2013	208	0.51 (0.24)	16.2 (1.72)
LO	2014	208	0.24 (0.13)	17.52 (1.69)
STBC	2010	774	0.04 (0.09)	10.28 (2.03)
STBC	2011	774	0 (0.01)	NA
STBC	2012	774	0.07 (0.08)	10.62 (2.00)
STBC	2013	774	0.02 (0.05)	9.34 (1.16)
STBC	2014	774	0.04 (0.08)	8.03 (1.19)
UC	2010	262	0.01 (0.03)	NA
UC	2011	262	0 (0.01)	NA
UC	2012	262	0.04 (0.14)	13.99 (2.81)
UC	2013	262	0.04 (0.12)	13.97 (2.47)
UC	2014	262	NA	12.91 (2.51)
UO	2010	1977	0.11 (0.14)	16.39 (2.46)
UO	2011	1977	0.12 (0.13)	NA
UO	2012	1977	NA	17.1 (2.31)
UO	2013	1977	0.23 (0.21)	15.96 (2.40)
UO	2014	1977	0.08 (0.10)	16.1 (2.53)
WC	2010	1822	0.1 (0.10)	11.19 (2.17)
WC	2011	1822	0.08 (0.13)	NA
WC	2012	1822	0.09 (0.09)	13 (2.09)
WC	2013	1822	0.15 (0.18)	12.11 (1.71)
WC	2014	1822	0.13 (0.16)	12.96 (2.23)
WN	2010	8062	0.35 (0.17)	NA
WN	2011	8062	0.27 (0.28)	NA
WN	2012	8062	NA	15.13 (2.14)
WN	2013	8062	0.48 (0.27)	14.03 (2.25)
WN	2014	8062	0.36 (0.20)	13.31 (1.56)

Sd= Standard deviations

Table A3: Evanno method on the STRUCTURE analysis of population subdivision in four brook trout stream drainages from Cape Race, Newfoundland.

Drainage	K	Mean LnP(K)	Stdev LnP(K)	ΔK
Coquita	1	-2658	0.05	NA
Coquita	**2**	**-2443**	**0.63**	**310.56**
Coquita	3	-2424	3.21	2.20
Coquita	4	-2398	7.83	2.36
Coquita	5	-2390	2.09	NA
O'Beck	1	-13317	0.08	NA
O'Beck	2	-13112	1.15	74.71
O'Beck	**3**	**-12993**	**2.01**	**3.36**
O'Beck	4	-12881	2.36	34.17
O'Beck	5	-12850	25.65	NA
Freshwater	1	-12786	0.20	NA
Freshwater	2	-12635	7.31	10.25
Freshwater	**3**	**-12559**	**4.28**	**39.87**
Freshwater	4	-12654	44.31	0.29
Freshwater	5	-12736	44.00	NA
Watern	1	-12018	0.10	NA
Watern	**2**	**-11826**	**2.91**	**89.89**
Watern	3	-11895	15.24	1.65
Watern	4	-11990	13.86	5.97
Watern	5	-12002	59.24	NA

With K the number of clusters, Mean LnP(K) the mean likelihood of K among runs, Stdev.LnP(K) the standard deviation of lmP(K) among runs, and ΔK the change in LnP(K). We highlighted in bold the inferred population clustering according to a combination of the ΔK, the ln P(K) values and by assessing the strength of individual assignments within clusters. In one drainage (O'Beck), the strength of individual assignments within clusters and isolation of the stream beds suggested that clustering was better when K=3 than when K=2.

Table A4: Test statistics for the final model terms predicting N_b in 12 brook trout populations from Cape Race, Newfoundland, Canada.

Parameter	df	F value	p value
ln (stream size)	1,10.7	26. 9	<0.001

Table A5: Mixed-effect modeling of the relationship between *N* and stream length among 12 brook trout populations from Cape Race, Newfoundland, Canada

Fixed effects estimates (SE)	
Intercept	6.59 (0.25)
ln (stream length)	1.28 (0.24)
Random effects variance components (Sd)	
Population	0.70 (0.84)
residual	0.07 (0.27)
R^2_m	0.65
R^2_c	0.97
n	63

SE=standard error; Sd=standard deviation; R^2_m = marginal R square (fixed effects only); R^2_c = conditional R square (fixed and random effects); n=number of observations

Table A6: Temporal variability in N, N_b and N_b/N within 12 brook trout populations from Cape Race, Newfoundland.

	Temporal CVs		
Pop	N	N_b	N_b/N
BC	0.191	0.343	0.468
CC	0.224	0.182	0.433
DY	0.325	0.779	0.709
HM	0.182	0.577	0.300
LB	0.193	0.354	0.358
LC	0.523	0.704	0.908
LO	0.212	0.910	0.929
STBC	0.284	0.576	0.853
UC	0.178	0.288	0.259
UO	0.265	0.329	0.408
WC	0.285	0.311	0.181
WN	0.193	0.318	0.530

Temporal coefficient of variations (CV) are measures of dispersion where CV=σ/μ (σ =standard deviation of N_b or N_b/N within each population and μ= mean N_b or N_b/N within each population).

Table A7: Spatial variability in N, N_b and N_b/N among 12 brook trout populations from Cape Race, Newfoundland.

	Spatial CV		
Year	N	N_b	N_b/N
2008	NA	1.055	NA
2009	NA	1.145	NA
2010	1.239	0.946	0.945
2011	1.196	1.646	1.348
2012	1.206	1.331	1.518
2013	1.585	1.004	0.994
2014	1.346	1.147	1.212
2015	1.167	NA	NA

Spatial coefficient of variations (CV) are measures of dispersion where CV=σ/μ (σ =standard deviation of N or N_b/N among populations for each year and μ= mean N_b or N_b/N for each year).

Table A8: Mixed-effect modeling of the relationship between incorrectly linked *N* and N_b, sampled from the same season (based on adult tagging for *N* estimation and juvenile YOY sampling for N_b estimation)

Fixed effects estimates (SE)		
	Intercept	3.79 (0.15)
	N	0.90(0.14)
Random effects variance components (Sd)		
	Population	0.17 (0.41
	Residual	0.34 (0.59)
R^2_m		0.59
R^2_c		0.73
n		53

SE= standard error; Sd= standard deviation; R^2_m = marginal R square (fixed effects only); R^2_c = conditional R square (fixed and random effects); n=number of observations; significance codes according to likelihood ratio test: 0***, 0.01 **, 0.01 *Significance level ***, p<0.001

Fig.A1: Relationship between *N* and stream size from a mixed model.

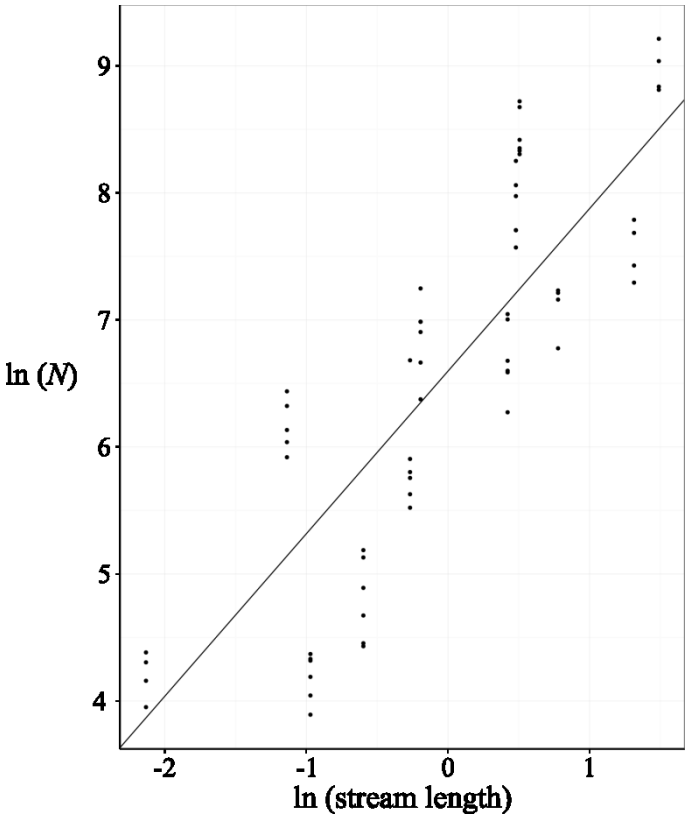
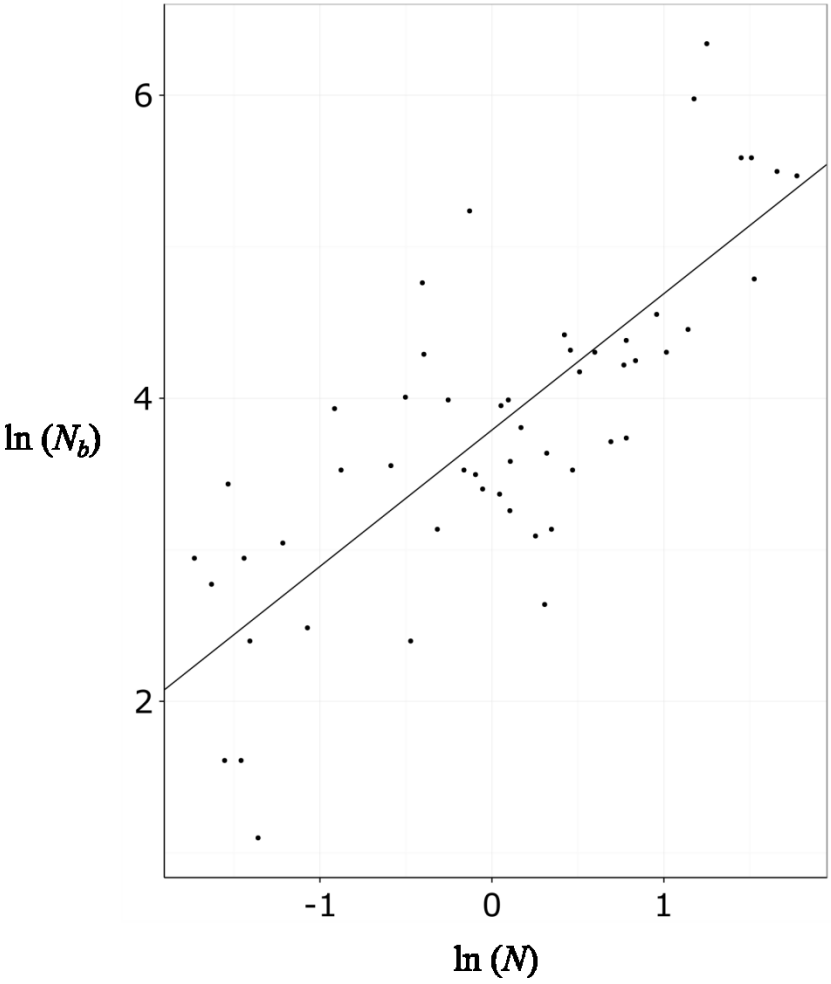


Fig.A2: Relationship between incorrectly linked N and N_b from a mixed model, sampled from the same season (based on adult tagging for N estimation and juvenile YOY sampling for N_b estimation).



Appendix B

Table B1: Spawning area and stream length in 9 brook trout populations from Cape Race, Newfoundland.

Fig. B1: The relationship between area available for spawning and stream length in 8 brook trout populations from Cape Race, Newfoundland.

The relationship between spawning habitat area and stream length in 8 brook trout populations from Cape Race, Newfoundland, Canada.

As a supplementary analysis, we investigated whether stream length and the area selected for spawning by brook trout within each stream were positively correlated. Indeed, spawning areas are usually associated with areas of upwelling groundwater with distinct chemical and physical properties (Curry & Neakes 1995), and natural brook trout populations are highly dependent on their availability and quality. Additionally, we found a strong, positive correlation between N_b and stream length. Thus, a positive correlation between spawning habitat area and stream length might suggest that spawning site availability may limit the number of breeding adults.

Methods

Cape Race brook trout aggregates were observed in nine populations during the spawning season (October) in 2010, 2011 and 2014. Spawning area was then calculated using a combination of Google Earth and Earthpoint (Table B1; http://www.earthpoint.us/Shapes.aspx). We used a linear mixed model with spawning area as a response variable and stream length as an explanatory variable.

Results

At first, we found a statistically significant positive correlation between stream length and area available for spawning ($p=0.02$, $df=7$, adj.$r^2=0.53$) but the relationship was mostly driven by a single outlier. We removed the outlier since its presence was driving non-normality even after log transformation; after removal of the outlier, we found no significant relationship between the area available for spawning and stream length (Fig. B1; $p=0.41$, $df=6$), which suggested that spawning site availability may not limit N_b in Cape Race streams. Of course, an analysis including more data points is necessary to better understand the relationship between spawning area and stream length.

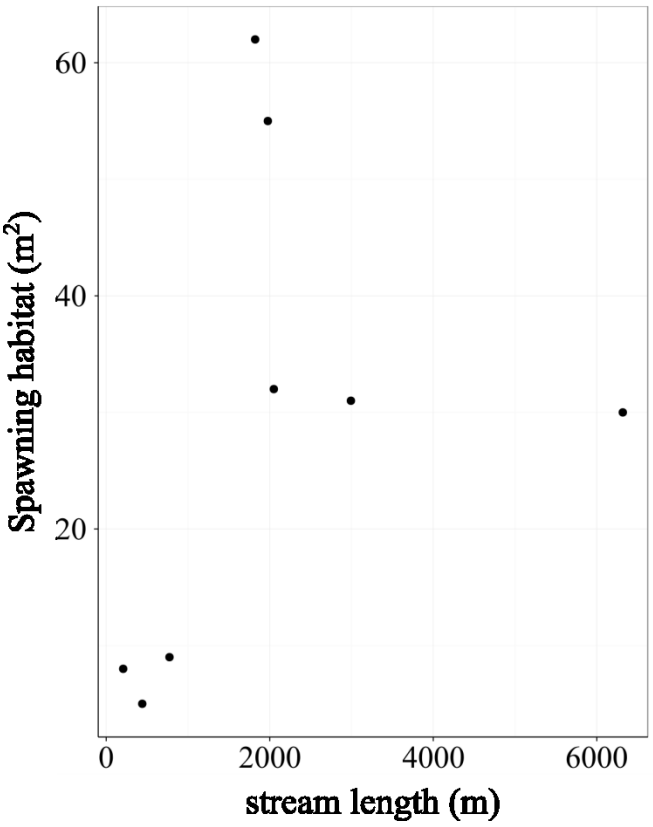
Reference

Curry RA, Noakes DLG, Curry RA, Neakes DLG (1995) Groundwater and the selection of spawning sites by brook trout (Salvelinus fontinalis). *Canadian Journal of Fisheries and Aquatic Sciences*, **52**, 1733–1740.

Table B1: Spawning area and stream length in 9 brook trout populations from Cape Race, Newfoundland.

Population	Total spawning area (m2)	Stream length (m)
BC	32	2050
STBC	9	774
WC	62	1822
WN	158	8062
CC	30	6318
LO	8	208
UO	55	1977
DY	5	441
LB	31	2993

Fig. B1: The relationship between area available for spawning and stream length in 8 brook trout populations from Cape Race, Newfoundland.



Appendix C: Forecasting N_b, N and stream length from the other in brook trout

Table C1: Results from the Model II regression of pairwise N_b, N and stream length. Results are also given from linear mixed model (LMM) for comparison.

Fig. C1: Model II regressions showing the relationships between N_b and N, N_b and stream length and N and stream length. Confidence intervals are indicated with dotted lines.

Forecasting N_b, N and stream length from the other in brook trout

To estimate the functional equations allowing forecasting one variable from another, we preferred model II regressions to standard regression techniques. Indeed, all variables measured in the experiment included random error (they were measured in different populations) as well as measurement error. Model II regression techniques also present the advantage that the slope estimators are symmetrical (the slope of N_b=f(N) is the reciprocal of the slope of N=f(N_b)). Several approaches to model II regression techniques can be used; here, we used major axis regression (MA) as it was reasonable to assume that the error variance of the variables were comparable since they were natural log-transformed (Jolicoeur 1990).

Regression analyses were performed using the statistical package lmodel2 (Legendre 2014) in R. The 95% confidence intervals for both slope and intercepts are provided from the MA regressions as well as from mixed model analyses for comparison. The linear mixed-model analyses were similar to the ones conducted previously, with the exception that here variables were not standardized to allow the generalization of the relationship to other brook trout systems. All variables were natural log-transformed.

Although we base our functional equation inferences on model II regression techniques, it is apparent from the results that the method of regression used (mixed models vs. model II regression) makes little difference with regards to the conclusion. This is a likely consequence of the strong correlation existing between the variables.

Results

As previously, model II regressions showed a significant positive correlation between N_b and N, N_b and stream length, and N and stream length (Table 1, Fig.1). Interestingly, the slope of the relationships were significantly different from 1, especially for the relationship between N_b and N. This suggests that N_b does not proportionally increase with N, thus providing further support for genetic compensation.

According to the MA regressions, the slope of the equation linking N_b to N was 0.68 where $\ln(N_b)=0.6760*\ln(N) - 0.6256$, with $r^2=0.56$; reciprocally, the slope of the equation linking N to

Nb was 1.48 where ln(N)=1.4794*ln(N_b)+0.9254 (r^2=0.56 as well). Additionally, the slope of the equation linking N to stream length was 1.19 where ln(N)=1.1903*ln (stream length) -1.6228, with r^2=0.70. Finally, the slope of the equation linking N_b to stream length was 0.73 where ln(N_b)=0.7309*ln(stream length) -1.2535 where r^2=0.47.

The slope and intercept parameters obtained from the mixed model analyses overlapped with those estimated from the model II regressions, implying that the results were comparable between the methods of analyses. However, parameters derived from the mixed model analysis were consistently lower than those derived from the regression II, which is not surprising since OLS regresisons tend to be biased low due to the presence of random errors (reviewed in Legendre and Legendre 1995).

References

Jolicoeur, P. (1990). Bivariate allometry: interval estimation of the slopes of the ordinary and standardized normal major axes and structural relationship. *Journal of Theoretical Biology*, 144(2), 275-285.

Legendre P, Legendre L (1988) *Numerical Ecology*.

Table C1: Results from the Model II regression of pairwise *Nb*, *N* and stream length. Results are also given from linear mixed model (LMM) for comparison.

Response variable	ln *Nb*		ln *N*
Explanatory variable	Ln *N*	ln stream length	ln stream length
slope (95% CI)	0.68 (0.50; 0.89)	0.73 (0.55; 0.95)	1.19 (1.01-1.41)
intercept (95% CI)	-0.63 (-2.02; 0.54)	-1.25 (-2.81; 0.03)	-1.62 (-3.14; -0.38)
r^2	0.56	0.47	0.7
LMM slope (95% CI)	0.56 (0.36; 0.75)	0.56 (0.27-0.86)	0.92 (0.58;1.25)
LMM intercept (95% CI)	0.17 (-1.15; 1.51)	-0.07 (-2.12, 1.98)	0.24 (-2.12; 2.59)
n	43	63	63

CI: confidence intervals, n: sample size.

Fig.C1: Model II regressions showing the relationships between N_b and N, N_b and stream length and N and stream length. Confidence intervals are indicated with dotted lines.